12

CRM 86-187 / August 1986

RESEARCH MEMORANDUM

CONFIRMING DIFFERENCES IN RELATIVE-VALUE PROFICIENCY MARKS

Paul W. Mayberry

A Division of CNA Hudson Institute

CENTER·FOR·NAVAL·ANALYSES

4401 Ford Avenue • Post Office Box 16268 • Alexandria, Virginia 22302-0268 • (703) 824-2000

Work conducted under contract N00014-83-C-0725.

SECURITY CLASSIFICATION OF THIS PAGE

AD-A175362

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION Unclassified	1b. RESTRICTIVE MARKINGS
2a. SECURITY CLASSIFICATION AUTHORITY	3. DISTRIBUTION / AVAILABILITY OF REPORT Approved for Public Release; Distribution Unlimited.
2b. DECLASSIFICATION / DOWNGRADING SCHEDULE	
4. PERFORMING ORGANIZATION REPORT NUMBER(S) CRM 86-187	5. MONITORING ORGANIZATION REPORT NUMBER(S)
6a. NAME OF PERFORMING ORGANIZATION Center for Naval Analyses 6b. OFFICE SYMBOL (If applicable) CNA	7a. NAME OF MONITORING ORGANIZATION Commandant of the Marine Corps (Code RDS)
6c. ADDRESS (City, State, and ZIP Code) 4401 Ford Avenue Alexandria, Virginia 22302-0268	7b. ADDRESS (City, State, and ZIP Code) Headquarters, Marine Corps Washington, D.C. 20380
8a. NAME OF FUNDING / ORGANIZATION Office of Naval Research 8b. OFFICE SYMBOL (If applicable) ONR	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER N00014-83-C-0725
8c. ADDRESS (City, State, and ZIP Code) 800 North Quincy Street Arlington, Virginia 22217	10. SOURCE OF FUNDING NUMBERS PROGRAM ELEMENT NO. 65153M; PROJECT NO. C0031; TASK NO.; WORK UNIT ACCESSION NO.

11. TITLE (Include Security Classification)

Confirming Differences in Relative-Value Proficiency Marks

12. PERSONAL AUTHOR(S)
Paul W. Mayberry

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Year, Month, Day)	15. PAGE COUNT
Final	FROM TO	August 1986	30

16. SUPPLEMENTARY NOTATION

17. COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB-GROUP	Comparison, Enlisted personnel, Evaluation, Hands-on tests, Job performance, Marine Corps personnel, MOS (Military Occupationnal Specialty), Performance (human), Performance tests, Productivity, Proficiency, Skills, Surveys
05	09		
05	10		

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

Knowing the magnitude of performance differences among individuals is critical in making personnel decisions in the Marine Corps. Recent research at the Center for Naval Analyses has focused on quantifying the magnitude of these performance differences. The purpose of this analysis was to cross-validate those results using hands-on measures of job performance and other published research findings.

20. DISTRIBUTION / AVAILABILITY OF ABSTRACT	21. ABSTRACT SECURITY CLASSIFICATION	
☐ UNCLASSIFIED/UNLIMITED ☒ SAME AS RPT ☐ DTIC USERS	Unclassified	
22a. NAME OF RESPONSIBLE INDIVIDUAL Lt.Col. G. W. Russell	22b. TELEPHONE (Include Area Code) (202) 694-3491	22c. OFFICE SYMBOL RDS-40

DD FORM 1473, 84 MAR 83 APR edition may be used until exhausted.

All other editions are obsolete.

SECURITY CLASSIFICATION OF THIS PAGE

A Division of CNA Hudson Institute

CENTER·FOR·NAVAL·ANALYSES

4401 Ford Avenue • Post Office Box 16268 • Alexandria, Virginia 22302-0268 • (703) 824-2000

3 September 1986

MEMORANDUM FOR DISTRIBUTION LIST

Subj: Center for Naval Analyses Research Memorandum 86-187

Encl: (1) CNA Research Memorandum 86-187, "Confirming Differences in Relative-Value Proficiency Marks," by Paul W. Mayberry, Aug 1986

1. Enclosure (1) is forwarded as a matter of possible interest.

2. Individuals differ in the quality of their performance within the Marine Corps. Accurate quantification of the magnitude of these performance differences has important implications for Marine Corps personnel decisions. This Research Memorandum examines Marine Corps job performance data and other research on industrial productivity to validate the relative value associated with differences in the Marine Corps proficiency marks.

William H. Sims
Director
Marine Corps Manpower
and Training Program

Distribution List:
Reverse Page

Subj: Center for Naval Analyses Research Memorandum 86-187

Distribution List

SNDL

A1	ASSTSECNAV MRA
A1	DASN - MANPOWER (2 copies)
A6	HQMC MPR
	Attn: Deputy Chief of Staff for Manpower (2 copies)
	Attn: Director, Personnel Procurement Division (2 copies)
	Attn: Director, Manpower Plans and Policy Division (2 copies)
	Attn: Director, Personnel Management Division (2 copies)
A6	HQMC TRNG (2 copies)
A6	HQMC RD&S (2 copies)
A6	HQMC RA (2 copies)
A6	HQMC AVN (2 copies)
E3D1	CNR
E3D5	NAVPERSRANDCEN
	Attn: Director, Manpower and Personnel Laboratory
	Attn: Technical Library
FF38	USNA
	Attn: Nimitz Library
FF42	NAVPGSCOL
FF44	NAVWARCOL
FJA1	COMNAVMILPERSCOM
FJB1	COMNAVCRUITCOM
FT1	CNET
V12	CG MCDEC

OPNAV

OP-01
OP-11
OP-12
OP-13

Other

Joint Service Job Performance Measurement Working Group (11 copies)

CRM 86-187 / August 1986

CONFIRMING DIFFERENCES IN RELATIVE-VALUE PROFICIENCY MARKS

Paul W. Mayberry

Marine Corps Operations Analysis Group

A Division of CNA *Hudson Institute*

CENTER·FOR·NAVAL·ANALYSES

4401 Ford Avenue • Post Office Box 16268 • Alexandria, Virginia 22302-0268

ABSTRACT

Knowing the magnitude of performance differences among individuals is critical in making personnel decisions in the Marine Corps. Recent research at the Center for Naval Analyses has focused on quantifying the magnitude of these performance differences. The purpose of this analysis was to cross-validate those results using hands-on measures of job performance and other published research findings.

EXECUTIVE SUMMARY

BACKGROUND

Knowing the magnitude of performance differences among individuals in terms of incremental value to the Marine Corps is critical in making personnel decisions. Given the current performance appraisal system of the Marine Corps, it was known that different personnel subgroups generally receive different performance ratings. However, the translation of these performance differences into incremental value to the Marine Corps had not been addressed.

Recent research by analysts at the Center for Naval Analyses [1] has focused on developing a methodology to quantify the magnitude of these performance differences. All Marines in ranks private through corporal (E1–E4) regularly receive performance evaluations from their supervising officers. These evaluations are quantified in terms of proficiency marks. Responding to a specially designed survey, Marine Corps officers were asked to translate their personal interpretations of the proficiency marks into estimates of increased value to the Marine Corps. The translation was then applied to existing proficiency marks for a large sample of first-term Marines. The purpose of this study was to cross-validate the magnitude of performance differences associated with proficiency marks as found by the survey.

FINDINGS

The analysis was based upon two data sources: a CNA study of job performance measurement and relevant research on industrial productivity. Table I summarizes the findings of the magnitude of individual performance differences from these two data sources as compared with the survey results. The magnitude of individual differences were expressed in two forms: the ratio of the standard deviation (SD) to the mean and the percent difference in performance between a group of top-level performers (95th percentile) and a group of bottom-level performers (5th percentile).

TABLE I

SUMMARY OF INDIVIDUAL PERFORMANCE DIFFERENCES

Performance indicator	Ratio of SD to mean	Difference between 95th and 5th percentile
Relative-value proficiency marks	29%	161%
Hands-on job performance tests	23%	127%
Industrial productivity	20%	106%

The survey results of relative-value proficiency marks are somewhat higher than the differences found for industrial productivity, but they are comparable to the Marine Corps job performance tests. Such findings are positive, in that the proficiency marks and hands-on performance tests were specifically designed to assess performance in the Marine Corps. Therefore, it is encouraging that the magnitude of the proficiency-mark differences resembles the magnitude of job performance test-score differences. The relatively large magnitude of the survey results may possibly be attributed to method variance. That is, the hands-on test scores and quantity of industrial output are essentially objective measures of performance. On the other hand, surveys are more subjective and ask persons to quantify concepts that they generally do not perceive to be quantifiable. Accordingly, the survey responses differ from the objective measures and, in this case, the difference is in the direction of overestimation.

Because of the design of the survey, caution must be exercised in interpreting the relative value associated with proficiency marks less than 4.0. The comparison of marks less than 4.0 is ambiguous because the exact reference point for comparison is unknown. However, given that few individuals receive marks less than 4.0, this caveat *does not apply* to a significant proportion of the proficiency-mark distribution.

CONCLUSION

The magnitude of performance differences found in the survey are comparable to the results noted for hands-on tests of job performance and similar to the findings of research on industrial productivity. This confirmation of the survey results is encouraging for the application of relative-value proficiency marks to manning issues within the Marine Corps.

TABLE OF CONTENTS

	Page
List of Tables	ix
Introduction	1
Individual Differences in Proficiency Marks	1
Individual Differences in Job Performance	5
Individual Differences in Productivity	8
Discussion	11
Conclusion	14
References	15
Appendix: Assessment of the "Value" of Manpower Quality Survey	A-1–A-2

LIST OF TABLES

		Page
1	Translating Proficiency Marks Into Relative-Value Estimates	3
2	Individual Differences in Job Performance for Three Marine Corps Specialties	7
3	Individual Productivity Differences	10
4	Summary of Individual Performance Differences	12

INTRODUCTION

Judgments of individual differences in performance are an inherent requisite in adequately manning the Marine Corps. Personnel managers are continuously confronted with the necessity to make relative judgments of performance differences among various personnel classifications (e.g., AFQT category) and groups (e.g., race, sex, educational level) in order to establish enlistment quotas. Generally, the managers operate under the assumption that "more is better," but they are constrained by the real-world conditions of insufficient resources, underqualified personnel, and policy requirements. Therefore, the central issue becomes one of operating within the worldly constraints to recruit the mix of personnel that will result in maximum performance for the Marine Corps.

The specification of the optimal personnel mix is directly related to the degree of various subpopulation performance differences. Although there is widespread evidence of performance differences across certain personnel subgroups, little research has directly addressed the issue of its magnitude. Recent work by May [1] has focused on quantifying the magnitude of these performance differences among various personnel categories. Based upon the operational performance appraisal system currently used by the Marine Corps, the study centered on determining the degree of relative value between adjacent marks on the performance rating scale.

INDIVIDUAL DIFFERENCES IN PROFICIENCY MARKS

The Marine Corps performance appraisal system requires supervisors to rate the performance of enlisted personnel. Each Marine in pay grades E1 through E4 receives two scores: a proficiency mark (indicates how well the Marine performed his or her primary job responsibilities) and a conduct mark (reflects the Marine's observance of laws and regulations, as well as positive contributions to the Marine Corps). Only the proficiency marks are considered in this study, as they are more closely associated with the

issue of job performance and individual differences. Proficiency marks are based on a 0-to-5 scale, with the true satisfactory range of performance being between 4 and 5.

May [1] surveyed a sample of Marine Corps officers and asked them to translate their personal performance rating systems into estimates of incremental value to the Marine Corps. In other words, how much more valuable to the Marine Corps is an individual who received a mark of 4.6 versus an individual who received a mark of 4.5? Table 1 summarizes the results of this survey given to 218 Marine Corps officers and shows the distribution of average proficiency marks for the FY1981 – FY1982 time period of Marines who enlisted in FY1981. A copy of the survey is provided in the appendix.

The survey asked for value comparisons between one-tenth-point intervals for marks between 4.0 and 5.0, and one-point intervals between 0.0 and 3.9 (e.g., 3.0 to 3.9 compared to 4.0). The average increase in value to the Marine Corps between adjacent rating grades ranged from 32.9 to 9.3 percent. Standard deviations (SDs) of these comparisons were of the same magnitude as their means, except for the comparisons below a proficiency mark of 4.0. Thus, there was much less agreement among the officers in trying to translate their personal evaluation systems into percent differences for those comparisons involving marks less than 4.0. This may have been a function of the survey design, as the requested comparisons below 4.0 are more ambiguous and are based on a one-point interval. That is, what value comparison is to be made between a mark of 4.0 and the range of 3.0 to 3.9 – 3.5 versus 4.0, 3.0 versus 4.0, 3.9 versus 4.0, or some "average" rating versus 4.0?

The percent differences in value for the comparisons of adjacent proficiency grades (e.g., the 4.6 to 4.7 comparison) from table 1 were converted into relative-value estimates for each proficiency grade (e.g., 4.6 and 4.7). This was accomplished by choosing an arbitrary reference point (in this case, 0.5) and assigning it a value of 1. Successively larger proficiency marks were incremented in proportion to the percent differences reported by the Marine Corps officers for each of the proficiency-mark comparisons.

TABLE 1

TRANSLATING PROFICIENCY MARKS INTO
RELATIVE-VALUE ESTIMATES

Proficiency mark	Percent difference		Relative-value proficiency marks	FY1981 recruits	
	Mean	SD		Percent of total	Cumulative percent
5.0			9.09	0.0	100.0
	18.7	23.7			
4.9			7.66	0.9	100.0
	16.6	18.8			
4.8			6.57	4.6	99.1
	16.5	17.2			
4.7			5.64	11.4	94.5
	16.7	24.8			
4.6			4.83	18.5	83.1
	15.2	16.2			
4.5			4.19	21.1	64.7
	16.9	19.6			
4.4			3.59	15.9	43.6
	12.9	12.9			
4.3			3.18	9.9	27.6
	10.6	9.8			
4.2			2.87	5.4	17.8
	9.7	9.4			
4.1			2.62	3.3	12.3
	9.3	10.6			
4.0			2.40	2.3	9.1
	32.9	71.3			
3.5			1.80	5.9	6.7
	24.8	32.6			
2.5			1.44	0.7	0.8
	17.3	46.9			
1.5			1.23	0.1	0.1
	23.1	106.1			
0.5			1.00	0.0	0.0

Note: The percent and cumulative percent values are based on proficiency marks assigned to 28,092 Marines. The proficiency marks less than 4.0 represent midpoints of a one-point interval (i.e., 3.5 represents 3.0 to 3.9).

Relative-value estimates for each proficiency mark allow for the determination of incremental value of one proficiency mark compared to another.

The relative-value estimates were calculated for Marines who received proficiency marks during FY1981 to FY1982. The mean relative-value estimate was found to be 4.1 with a standard deviation of 1.2. However, caution must be exercised in the interpretation of these numbers because linear interpolations to a one-tenth-point interval were used for the comparisons that were measured on a one-point interval. Again, this points to the difficulty in interpreting the survey results for proficiency marks less than 4.1. Expressing the standard deviation of the relative-value estimates as a percentage of the mean shows that a one-standard-deviation change (i.e., 1.2 points on the relative-value estimate scale) will on average result in a 29-percent change in relative value.

Upon examining the full distribution of proficiency marks, a value of 3.8 was found to correspond to the 5th percentile, and a mark of 4.7 represented the 95th percentile. The relative-value estimates of these percentiles were 2.16 (determined by linear interpolation for the relative-value estimate between 3.5 and 4.0) and 5.64, respectively. The percent difference in the performance levels of these two groups was found by taking the ratio of the percentiles, dividing by the base 5th percentile, and multiplying by 100, that is, $\frac{5.64-2.16}{2.16} \times 100$. Accordingly, the percent difference between the high and low performer was found to be 161 percent. This implies that the high performers are 161 percent more valuable to the Marine Corps than their low-performing counterparts.[1]

[1]Other studies of individual performance differences often use the ratio of the 95th to 5th percentiles as the metric of difference, in contrast to percent difference. Both metrics yield the same result. However, it is important that the two metrics are not used interchangeably. While the percent difference between the groups is 161 percent (the high group is 161 percent more valuable than the lower group), the ratio of the 95th to 5th percentiles for the relative-value proficiency marks is 2.61 (the high group is 2.61 times more valuable than the lower group). A ratio of 2.61 and 161 percent have the same meaning. To prevent confusion, the percent difference metric will be used throughout this paper for comparing the high and low performance groups.

The results presented in table 1 and the two measures reflecting the magnitude of individual differences for the relative-value estimates are the cornerstone for any analyses incorporating the magnitude of differential performance by various personnel subgroups. Given the potential applications of this performance-difference methodology to a variety of Marine Corps manning issues, it is necessary that studies be conducted to substantiate the findings of these survey data. While such a charge appears to be straightforward, it is certainly ambitious, as evidenced by the dearth of work in this area.

Therefore, the purpose of this study is to use existing data sources in an effort to confirm the magnitude of performance differences obtained from the survey. Data from a 1981 feasibility study of job performance measurement conducted by the Marine Corps were used in estimating the magnitude performance differences, and these findings were compared to the survey data. Further analysis of these data, while not collected for this specific purpose, should be informative with respect to the question of the magnitude of performance differences. In addition, a pertinent article by Schmidt and Hunter [2] on individual differences in industrial productivity is reviewed and compared to the survey findings.

INDIVIDUAL DIFFERENCES IN JOB PERFORMANCE

In 1981, the Marine Corps conducted a study to determine the feasibility of measuring job performance using hands-on tests. Although there were problems noted with the project [3], the overall results were positive. A secondary goal of the project was to investigate the potential use of substitute measures of job performance that were correlated with the hands-on measures but would be more economical to administer. Proficiency marks were one such substitute measure that was studied.

Hands-on tests of job performance were developed for three military occupational specialties (MOSs):

- Ground Radio Repair: high technical requirements, 37 weeks of formal school training.
- Automotive Mechanic: moderate technical requirements, 13 weeks of formal school training.
- Infantry Rifleman: low technical requirements, 5 weeks of formal school training.

Only first-term enlistments were included in the study, but these personnel reasonably represented the full range of experience for a first-term Marine. The test administrators for each specialty were senior Marine Corps enlisted personnel with relevant job experience in their respective fields. The testing for all three specialties was conducted over a 3-month period.

The statistical summary of the scores resulting from the hands-on tests for these three specialties is presented in table 2. The score scales for the three tests differ significantly because each test was composed of differing numbers of tasks. The Automotive Mechanic scale is an efficiency score, which reflects units of performance per unit of time.

The score distributions for each test appear to be reasonable; that is, the distributions cover about one and a half standard deviations above and below the mean. However, the distribution for the Ground Radio Repair test is slightly negatively skewed, while the converse is true for the tests of the other two specialties.

In terms of individual differences with respect to the job performance measures, the ratios of the hands-on test standard deviation to the mean ranged from 20.5 percent to 26.9 percent. The largest percentage was noted for the Infantry Rifleman specialty. Results of the same ordering were found for the comparison of the 95th percentile to the 5th percentile. These percentages ranged from a low of 95 percent to upwards of 150 percent. To confirm these findings and to determine the effect the observed score scale

TABLE 2

INDIVIDUAL DIFFERENCES IN JOB PERFORMANCE
FOR THREE MARINE CORPS SPECIALTIES

Performance scores	Ground Radio Repair	Automotive Mechanic	Infantry Rifleman
N	135	173	259
SD	1.81	.102	37.1
Mean	7.99	.497	138.0
Ratio	22.7%	20.5%	26.9%
Range			
Minimum	3.14	.308	42.0
Maximum	10.00	.870	226.0
Observed			
95th percentile	9.92	.685	200.0
5th percentile	4.39	.352	80.0
Percent difference	126%	95%	150%
Normalized			
95th percentile	10.97	.665	199.1
5th percentile	5.01	.329	76.9
Percent difference	119%	102%	159%

Note: The ratio expresses the percent change of a standard deviation unit in the hands-on score mean. The percent difference is calculated for both the actual observed 95th percentile to observed 5th percentile and the normalized 95th percentile to normalized 5th percentile.

distribution had on the percentages, the percentile comparison was computed for the normalized score scale also. These would be the expected results if the hands-on test scores had been normally distributed. The results are in the expected direction; that is, for the slightly positively skewed distributions, Automotive Mechanic and Infantry Rifleman, the percentages increased compared to the observed score scale while the opposite was true for the negatively skewed Ground Radio Repair distribution. By normalizing the distributions, the process has a compensating effect to overcome the direction of the skewness, but the results were essentially the same as those of the original scales.

In summary, a one-standard-deviation change in the hands-on test scores is about 23 percent of the mean, averaging across all three MOSs. With respect to the comparison of high versus low performers for the normalized score scale, the percent difference is in the range of about 127 percent. By examining three MOSs, three independent replications were found, which is a strong confirmation for the noted magnitude of the performance differences and implies stability of the magnitude of individual performance differences across specialties.

INDIVIDUAL DIFFERENCES IN PRODUCTIVITY

To the extent that individuals differ in terms of their productivity or output and that these differences can be quantified, significant gains can be obtained in organizational productivity through improved selection systems, or significant savings can be realized in reduced personnel required to complete a fixed level of work. Schmidt and Hunter [2] specifically address this problem of quantifying individual differences for these two purposes through a process called validity generalization. Essentially the technique is a meta-analysis of past research findings, based on the premise that each finding is simply an observation from an entire distribution of possible findings. Therefore, analysis of a collection of studies provides a much more powerful estimate of the underlying population parameters than any single study.

Schmidt and Hunter examined a 50-year span of research for studies that provided the information to calculate the productivity percentages of individual differences. Their efforts revealed 18 sources that had investigated some 40 jobs. They separated the jobs into three clusters according to their compensation system: incentive or piecerate work, nonincentive or straight pay, and "uncertain" pay, so-called because the compensation process was not fully described in the study. It was hypothesized that the uncertain jobs were more like the straight-pay system than the piecerate work. Given that Marine Corps personnel are not compensated on a piecerate basis, only the nonincentive and uncertain job results are presented here.

Table 3 presents a summary of the results for 28 jobs examined for the nonincentive and uncertain compensation systems. The results are very similar for the two compensation systems, and the table also notes the combined results. The jobs that were specifically classified as nonincentive tended to be more homogeneous (i.e., the standard deviations for their 14 ratios were smaller) than those of the uncertain classification. Likewise, the magnitude of the percentages for the nonincentive jobs was slightly less than the percentages for the uncertain group. Given that the percentages for the incentive condition (which are not presented here) were even smaller than the nonincentive system, it appears that the uncertain jobs are more similar to the nonincentive jobs than to the incentive jobs. Thus, it would be justifiable to combine the nonincentive and uncertain groups. The overall results show that the standard deviation of productivity is approximately 20 percent of the mean. Comparing the top 5 percent of the workers versus the bottom 5 percent indicates that there is approximately a 106-percent difference in their output.

Despite the power gained by analyzing a variety of studies, two factors should be noted that may restrict the magnitude of the noted productivity differences. First, the findings could be conservative because the studies examined by Schmidt and Hunter included only experienced job incumbents – "A number of studies presented findings separately for experienced employees and all employees; in such cases, only the results for the expe-

TABLE 3

INDIVIDUAL PRODUCTIVITY DIFFERENCES

Compensation system	Number of jobs	Average ratio of SD to mean	Average difference
Nonincentive	14	18.5%	93%
Uncertain	14	21.5%	120%
Combination	28	20.0%	106%

Note: The combination compensation system represents the average of the 28 jobs for both the nonincentive and uncertain systems. The average percent difference is based on the comparison of the 95th to the 5th percentile. This table is a summary of the results presented by Schmidt and Hunter [2].

rienced employees were used. Other studies specified that only data from experienced employees were analyzed" [2, page 408]. If the studies had incorporated less experienced personnel or applicants, the standard deviation could disproportionately increase relative to the decrease in the mean productivity output, and thus the ratio of standard deviation to mean would be higher. Certainly, the percent difference between the 95th and 5th percentiles would have increased given this more diverse range of individual productivity. Second, these results may be an understatement of the true productivity differences between individuals because only quantity of output was considered with no concern for the quality of that output. If quantity and quality are correlated to any degree, that is, high producers of output are also high-quality producers, then the results of table 3 clearly underestimate the degree of individual performance differences. Likewise, depending upon the severity of the errors committed by less productive personnel, the percentages could significantly differ with the high producers being even more productive than their less productive counterparts.

DISCUSSION

Table 4 summarizes the findings of the three studies of individual performance differences: May's investigation of relative-value proficiency marks, the job performance study of hands-on tests for three Marine Corps MOSs, and Schmidt and Hunter's meta-analysis of industrial productivity.

For both measures of performance differences, the survey results of relative-value proficiency marks are somewhat higher than the differences found for industrial productivity, but they are comparable to the Marine Corps job performance tests. Such findings are positive, in that the proficiency marks and hands-on performance tests were specifically designed to assess performance in the Marine Corps. Therefore, it is encouraging that the magnitude of the proficiency-mark differences resembles the magnitude of job performance test-score differences. The relatively large magnitude of the survey results may possibly be attributed to method variance. That is, the hands-on test scores and quantity of industrial output are essentially objective measures of performance. On the other hand, surveys are more

TABLE 4

SUMMARY OF INDIVIDUAL PERFORMANCE DIFFERENCES

Performance indicator	Ratio of SD to mean	Difference between 95th and 5th percentiles
Relative-value proficiency marks	29%	161%
Hands-on job performance tests	23%	127%
Industrial productivity	20%	106%

subjective and ask persons to quantify concepts that they generally do not perceive to be quantifiable. Accordingly, the survey responses differ from the other more objective measures, and in this case, the difference is in the direction of overestimation.

Recall that the percentages reported for industrial productivity were thought to be conservative for two reasons. First, the industrial results were based on only experienced job incumbents. This effect is confirmed with the job performance tests in that the Infantry Riflemen were found to have the largest performance differences. This specialty certainly has the largest diversity of experience and ability compared to Ground Radio Repair and Automotive Mechanic, which have higher entrance standards and technical training requirements. Second, the industrial performance differences were based merely on quantity of output, with no regard to product quality or severity of production errors. The hands-on tests, while not complete representations of the vast array of job requirements, were more extensive than the industrial output measure and included evaluations of the quality of performance according to standard operating procedures. Thus, the magnitude of the industrial productivity percentages may be an underestimation of the performance differences, and therefore, be more in line with the performance tests or relative-value proficiency marks.

A caveat must be reiterated concerning the application of the relative-value proficiency marks. That is, the relative value of a proficiency mark less than 4.0 is difficult to interpret given the design of the survey. The survey was ambiguous concerning the comparisons of marks less than 4.0 because comparison intervals of a full point were used as opposed to intervals of a tenth of a point. While it is understandable why this strategy was taken, it is unfortunate that the tenth-of-a-point intervals did not extend farther down the proficiency scale to include all proficiency marks that have true meaning (down to 3.5 would have included about 98 percent of all proficiency marks). Whether the officers were actually comparing the *midpoints* of the interval to the next higher midpoint or value is not known because no such instructions were given. Therefore, linear interpolation between the interval midpoints is questionable. Is 3.5 equally different from

3.6 as 3.9 is from 4.0? This would not be expected given the trend of increasing differences in table 1 for marks greater than 4.0. In addition, the selection of a specific reference point for the relative-value scale has some impact on the interpolation of relative-value marks to the one-tenth-point interval. That is, different relative-value proficiency marks are calculated for the one-tenth-point intervals if the interpolation is made between 3.0 and 4.0 (a reference point of 0.0) versus between 3.5 and 4.0 (a reference point of 0.5). For 3.0 compared to 4.0, the 32.9-percent increase in value (see table 1) is distributed over ten intervals as opposed to only five intervals for the comparison of 3.5 to 4.0. By using 3.5 as the comparison point, as was the case in May's work, the difference between the 95th and 5th percentiles is 14 percent higher than it would be if 3.0 were the comparison point. However, it is believed that the 3.5 value is the more realistic comparison point and the one most used by the Marine Corps officers in making their decisions. If any other comparison point were used that was less than 3.5, the percent differences would more closely approach the findings noted for the performance tests and industrial productivity.

CONCLUSION

The magnitude of performance differences found in the survey are comparable to the results noted for hands-on tests of job performance and similar to the findings of research on industrial productivity. This confirmation of the survey results is encouraging for the application of relative-value proficiency marks to manning issues within the Marine Corps.

REFERENCES

[1] CNA, Research Contribution 550, *Translating Supervisory Ratings Into Relative Value Estimates*, by Laurie J. May, Unclassified, July 1986

[2] Schmidt, F. L. and Hunter, J. E. "Individual Differences in Productivity: An Empirical Test of Estimates Derived From Studies of Selection Procedure Utility." *Journal of Applied Psychology*, 68 (Fall 1983): 407-414

[3] CNA, Research Contribution 540, *Examining the Validity of Hands-On Tests of Job Performance*, by Paul W. Mayberry, Unclassified, May 1986

APPENDIX

ASSESSMENT OF THE "VALUE" OF MANPOWER QUALITY SURVEY

APPENDIX

ASSESSMENT OF THE "VALUE" OF MANPOWER QUALITY SURVEY

HQMC is conducting research to develop a methodology for establishing Marine Corps manpower quality requirements. Quality is represented by such oategories as education credentials (HSG, NHSG, GED, etc.) and mental group (MG I thru V). At the present time, none of the Services use a DoD-recognized or accepted analytic procedure for determining these requirements. The current Marine Corps study effort is being performed by the Marine Corps Operations Analysis Group (MCOAG) of the Center for Naval Analyses (CNA). It involves use of the principles of economic analysis, including identification and measurement of the costs and benefits associated with accessing, training, and retaining differing quality Marines. In order to more appropriately assess the RELATIVE VALUE to the Marine Corps of different quality Marines, MCOAG requires assistance in the form of the professional judgment of Marine leaders.

Although proficiency and fitness report marks provide a RANK ORDER of performance, they do not provide an indication of HOW MUCH better the performance aotually is. That is, a commanding officer knows that a "4.8 Marine" is more valuable to the Marine Corps than a "4.2 Marine," but "how much more?" is not known. Since no universally acceptable substitute for performance (i.e., rifle range or PFT score) exists against which to measure and quantify the "general value to the Marine Corps" of different quality Marines, it is hoped that the professional judgment of Marine Corps leaders can be used to quantify the MAGNITUDE of this differenoe by considering the current performanoe evaluation systems. These estimates will be combined with those of all Marine leaders oompleting the form to produce an overall AGGREGATE table for translating performanoe evaluation marks into ESTIMATES OF RELATIVE VALUE to the Marine Corps.

Given the current political climate and the obvious Congressional interest in the DoD and Marine Corps budgets (active duty manpower represents 58% of the total USMC budget), development of an analytically sound methodology for projecting Marine Corps manpower quality requirements is of top priority. While many factors will go into the final choice of a proficiency or a fitness report mark given to an individual Marine, the interest in this assessment is how much more value to the Marine Corps does this mark represent compared to the mark on the grading scale just below it. Although certainly difficult, the judgmental effort requested in support of this research deserves your best shot.

Fill in the two tables below with the percent (%) that best reflects your estimate of the difference between adjacent scores. Proficiency marks are shown in the standard 5-point scale. Fitness report marks are shown for the "general value to the service" category. In between each pair of marks (going up the scale), write your estimate of the percent increase in VALUE TO THE MARINE CORPS between Marines receiving these marks. Estimates should represent your opinion of the incremental increase (%) over the next lower mark, and should be made independent of grade and MOS (i.e., consider that grade and MOS are identical in making your assessment of Marines being marked on the scoring system).

If, for example, in your experience, Marines to whom you historically have given proficiency marks of 4.5 are "worth 'X'% more to the Marine Corps" than those to whom you have given a mark of 4.4, then you should enter 'X' in the blank between 4.4 and 4.5. Or, if a Marine accomplishes twice as much as another, then the former is worth 100% more to the Marine Corps. Please complete all 23 blanks, and indicate the percent for each incremental increase that best reflects YOUR PERSONAL SCORING SYSTEM (note that the sum of your responses in each column DOES NOT have to total 100%):

PROFICIENCY MARKS

Scale	Percent Increase in "Value to the USMC"
5.0	
	)----------
4.9	
	)----------
4.8	
	)----------
4.7	
	)----------
4.6	
	)----------
4.5	
	)----------
4.4	
	)----------
4.3	
	)----------
4.2	
	)----------
4.1	
	)----------
4.0	
	)----------
3.0 - 3.9	
	)----------
2.0 - 2.9	
	)----------
1.0 - 1.9	
	)----------
0.0 - 0.9	

FITNESS REPORT MARKS

Scale	Percent Increase in "Value to the USMC"
OS	
	)----------
EX - OS	
	)----------
EX	
	)----------
AA - EX	
	)----------
AA	
	)----------
AV - AA	
	)----------
AV	
	)----------
BA - AV	
	)----------
BA	
	)----------
UN	

Your MOS_____